REPORT DOCUMENTATION PAGE		Form Approved OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE 31-08-2000	3. REPORT TYPE AND DATES COVERED FINAL: 01-05-1996 → 31-05-2000

4. TITLE AND SUBTITLE Adaptive Control and Navigation of Autonomous Mobile Robots	5. FUNDING NUMBERS N00014-96-1-0772

6. AUTHOR(S)
Paolo Gaudiano

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Boston University Dept. of Cognitive & Neural Systems 677 Beacon St. Boston, MA 02215	8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Office of Naval Research Program Officer: Teresa McMullen	10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release Distribution Unlimited	12b. DISTRIBUTION CODE Unlimited

13. ABSTRACT (Maximum 200 words)

This grant was used to support four different lines of research in the Neurobotics Lab at Boston University: Adaptive control of a mobile robot using unsupervised neural networks; Sensor Fusion for localization of a mobile robot; real-time visual tracking and positioning; sonar object recognition. All of these projects adhere the Neurobotics Lab's goal of using neural networks and other biomimetic approaches for sensory processing and control in mobile robotics.

14 20000905 107

15. NUMBER OF PAGES 15
16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT UNCL.	18. SECURITY CLASSIFICATION OF THIS PAGE UNCL.	19. SECURITY CLASSIFICATION OF ABSTRACT UNCL.	20. LIMITATION OF ABSTRACT

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102   UNCL (1)

DTIC QUALITY INSPECTED 4

Final Technical Report
Young Investigator Award, ONR-N00014-96-1-0772
Paolo Gaudiano, PI

This is the final technical report for the Office of Naval Research Young Investigator Award, titled "Adaptive control and navigation of autonomous mobile robots."

Final Period Report

The final year of this project was primarily spent winding down sponsored research activities, as the grant was scheduled to expire on 30-04-1999 but was prolonged to 31-05-2000 on a no-cost extension. Having graduated several PhD students who received support on this grant, the grant research in the period from 01-09-1999 through 31-05-2000 focused on a final project involving the use of sonar echoes for recognition of objects.

In collaboration with M. Ihsan Ecemis, who successfully completed and defended his Ph.D. dissertation this summer, we refined the sonar recognition system that was developed during earlier parts of the project.

From a scientific standpoint, we developed more sophisticated software for pre-processing and classification of sonar echoes, achieving very impressive results. The sonar is able to recognize one of several trained objects, showing robust recognition with multiple objects even at different distances and aspects.

As an attachment, we have enclosed a paper that will appear in the forthcoming conference "Sensor Fusion and Decentralized Control in Robotic Systems", part of SPIE's "Intelligent Systems for Advanced Manufacturing", to be held in Boston on 5-8 November, 2000. The paper describes some of the new methods we have devised to extract information from the echo which can be used for classification and recognition tasks.

The system is able to recognize objects that appear very similar: we trained the system to recognize three different faces at a fixed distance, and even with only several seconds of training on each face, the system was able to recognize the faces with an accuracy of over 70

The system's impressive performance has not gone unnoticed: in the final year of this grant we have received a grant from IS Robotics (Somerville, MA) to transfer some of this technology to ISR's mobile robotics platforms. Furthermore, a short segment on this work was featured on *Robocritters*, a documentary made by the British Broadcasting Corporation, which aired on the BBC with great success, and apparently also on the Learning Channel in the USA (neither the PI nor the Ph.D. student have the Learning Channel).

The following publications supported by this grant have appeared in the period 9/1/99–5/31/00:

- Ecemis, M.I. and Gaudiano, P. (2000). Object recognition system using sonar. Invited paper to appear in *SPIE Sensor Fusion and Decentralized Control in Robotic Systems*. Boston, MA, November 2000.

- Chang, C. and Gaudiano, P. (2000). Biomimetic Robotics.

- Ecemis, M. I. and Gaudiano, P. (1999). Object Recognition with Ultrasonic Sensors. *1999 IEEE International Symposium on Computational Intelligence in Robotics and Automation* (CIRA). Monterey, CA, 250–255.

- Sahin, E. and Gaudiano, P. (1999). Development of a Visual Object Localization Module for Mobile Robots. *Eurobot'99: Third European Workshop on Advanced Mobile Robots.* Zurich, Switzerland, September 1999.

Entire Performance Period

This has been a very productive grant during its four-year period. The funding was used to create the *Neurobotics Lab,* a laboratory for the application of neural networks and other biomimetic techniques to perception and control for mobile robots.

The Neurobotics Lab has generated a large number of publications and presentations, and, through collaborations with other academic units and with industry, it has trained several graduate students and achieved significant technology transfer.

Even though the PI has terminated his appointment at Boston University, the research developed under this grant has already had a significant impact that will undoubtedly outlast the original lab facility.

The research performed in the Neurobotics Lab has covered several topics that we now summarize.

Adaptive Control

With Dr. Carolina Chang, now a Professor at the Universidad Simon Bolivar in Caracas, Venezuela, we developed a neural network that could learn without supervision to control a robot's movement while avoiding obstacles and approaching lights.

The most significant innovation of this work was that our network requires no information about the robot's kinematics, or the nature and calibration of the robot's sensors. We were able to use exactly the same network to learn to control a Pioneer 1 (a medium-sized robot with five frontal and two lateral sonars) and a Khepera (a miniature robot with six frontal and two rear infrared proximity detectors). The ability to learn without supervision and without knowledge of the robot's kinematics and sensors makes our neural network very robust and flexible.

This project gave rise to some technology transfer: the PI received support of the *mediaCenter, GmbH,* an industry-sponsored science center located in Friedrichshafen, Germany, to apply some of the technology developed under this grant. Specifically, we developed a simulated home environment in which the miniature mobile robot Khepera could navigate while performing simple searching and surveillance tasks. All the low-level approach and avoidance behaviors on the Khepera were controlled using the neural networks described in earlier parts of this document. This project was highly successful as it led to the development of a unique simulated home environment (this was also the starting point of a project described later), and several useful technologies for real-time tracking and control of the mobile robot Khepera. This collaboration significantly enhanced the productivity of the Neurobotics Lab as a whole by providing funds for two graduate research assistants, two robots and three computers.

Sensor Fusion

The raw sensory input available to a mobile robot suffers from a variety of shortcomings. Sensor integration can yield a percept more veridical than is available from any single sensor input. In this project, the PI collaborated with Siegfried Martens, using the fuzzy ARTMAP neural network to integrate sonar and visual sonar on a B14 mobile robot. The neural network learned in a self-supervised fashion to associate specific sensory inputs with a corresponding distance metric. Once

trained, the network yielded predictions of range to obstacles that are more accurate than those provided by either sensor type alone. This improvement in accuracy was shown to hold across all distances and angles of approach tested.

Real-time Visual Tracking and Positioning

This project, in collaboration with Erol Sahin and Bob Wagner, extended the lab's previous work on visual tracking and localization originally developed for the simulated house environment. The work proceeded along two main lines. E. Sahin developed a system for testing visual navigation algorithms in real time with the Khepera mobile robot. The system combines information from on-board and overhead cameras to obtain real-time quantitative measurements of the accuracy of a localization algorithm. The algorithm makes it possible to visualize in real time on a computer screen the performance of any mapping and positioning algorithms, overlaying them on a live overhead camera image, thereby getting an immediate qualitative and quantitative measure of the algorithm's performance.

Using a similar tracking algorithm, Bob Wagner has collaborated over the course of three summers with Alan Schultz of the Naval Research Laboratory to develop a real-time tracking system for B-14 and Nomad robots with pan-tilt cameras. This work may be extended in the future to perform face tracking.

Sonar Object Recognition

The last project, which was already described in the first section of this report, was begun in collaboration with Bill Streilein. For the project has developed a biologically-inspired sonar sensor. It is known that bats and dolphins can use sonar for impressive recognition and localization tasks. We devised a way of recognizing objects using a neural network and a standard Polaroid sonar. With this system we were able to categorize a variety of typical indoor objects (chair, wall, box, trash can, etc.) with high accuracy independently of object distance and orientation.

Publications: Entire Period

1. Ecemis, M.I. and Gaudiano, P. (2000). Object recognition system using sonar. Invited paper to appear in *SPIE Sensor Fusion and Decentralized Control in Robotic Systems*. Boston, MA, November 2000.

2. Chang, C. and Gaudiano, P., Editors (2000). Special Issue on Biomimetic Robotics, **30**.

3. Ecemis, M. I. and Gaudiano, P. (1999). Object Recognition with Ultrasonic Sensors. *1999 IEEE International Symposium on Computational Intelligence in Robotics and Automation* (CIRA). Monterey, CA, 250–255.

4. Sahin, E. and Gaudiano, P. (1999). Development of a Visual Object Localization Module for Mobile Robots. *Eurobot'99: Third European Workshop on Advanced Mobile Robots*. Zurich, Switzerland, September 1999.

5. S. Martens, G. A. Carpenter and P. Gaudiano (1998). Neural network sensor fusion for spatial visualization. *SPIE Conference on Sensor Fusion and Decentralized Control in Robotic Systems*. **SPIE:3523**, 100–111.

6. E. Sahin and P. Gaudiano (1998). Mobile robot range sensing through visual looming. *Proceedings of the International Symposium on Computational Intelligence in Robotics and Automation (CIRA)*, 370–375. Gaithesburg, MD, September 1998.

7. S. Martens, P. Gaudiano and G. A. Carpenter (1998). Mobile robot sensor integration with fuzzy ARTMAP. *Proceedings of the International Symposium on Computational Intelligence in Robotics and Automation (CIRA)*, 307–312. Gaithesburg, MD, September 1998.

8. W. W. Streilein, P. Gaudiano and G. A. Carpenter (1998). A neural network for object recognition through sonar. *Proceedings of the International Symposium on Computational Intelligence in Robotics and Automation (CIRA)*, 271–276. Gaithesburg, MD, September 1998.

9. E. Sahin and P. Gaudiano (1998). Visual Looming as a range sensor for mobile robots. *Proceedings of the Fifth International Conference on Simulation of Adaptive Behavior*, 114–119. Zurich, Switzerland, August 1998.

10. Gaudiano P., Przybyszewski, A.W., van Wezel, R.J.A., van de Grind, W.A. (1998). Spatial Asymmetries in Cat Retinal Ganglion Cell Responses. *Biological Cybernetics*, **79**, 151–159.

11. Chang, C. and Gaudiano, P. (1998). Application of biological learning theories to mobile robot avoidance and approach behaviors. *Journal of Complex Systems*, **1**, 79–114.

12. P. Gaudiano and C. Chang (1997). Adaptive obstacle avoidance with a neural network for operant conditioning: experiments with real robots. *IEEE International Symposium on Computational Intelligence in Robotics and Automation*, 13–18. Monterey, CA, July 1997.

13. C. Chang and P. Gaudiano (1997). A neural network model of avoidance and approach behaviors for mobile robots. *World Multiconference on Systemics, Cybernetics and Informatics*. Caracas, Venezuela, July 1997.

14. Chang C. and Gaudiano P. (1997). Neural competitive maps for reactive and adaptive navigation. *Proceedings of the 2nd International Conference on Computational Intelligence and Neuroscience*, 19-23. Research Triangle Park, North Carolina, March 1997.

15. Gaudiano P. (1997). Being in the Right Place at the Right Time. *Synthese*, **112**, 125–134.

16. Gaudiano P., Guenther F.H., Zalama, E. (1997). The Neural Dynamics Approach to Sensory-Motor Control: Overview and Recent Applications in Mobile Robot Control and Speech Production. In van der Smagt, P. and Omidvar, O. (Eds.), *Neural Systems for Robotics*. San Diego: Academic Press. Ch.6, 153–194.

PI Presentations: Entire Period

1. March 5, 1999: "Neurobotics Lab research: Learning, vision and sonar recognition with mobile robots." Invited talk, *Fourth International Workshop on Neural Networks in Applications*, Magdeburg University, Germany.

2. March 3, 1999: "Neurobotics Lab research: application of learning theories to mobile robot navigation." Invited colloquium, University of Ilmenau, Germany.

3. July 10, 1998: "An overview of research in the Neurobotics Lab." Invited colloquium, Navy Center for Applied Research in Artificial Intelligence, Naval Research Laboratory, Alexandria, VA.

4. May 7, 1998: "An overview of the Neural Dynamics Approach." Invited colloquium, Center for Neuro-Mimetic Systems, Swiss Federal Institute of Technology, Lausanne, Switzerland.

5. April 29, 1998: "Sensor fusion for robot localization with an ARTMAP neural network." Invited seminar, Department of Neural Information Processing, University of Ulm, Germany.

6. April 27, 1998: "The importance of sensory-motor coordination: some examples from the Neurobotics Lab." Brown Bag Colloquium Series, Artificial Intelligence Laboratory, University of Zurich, Switzerland.

7. April 24, 1998: "An overview of research in the Neurobotics Lab: mobile robot applications in vision, localization and sonar object recognition." Invited seminar, Institute of Robotics and System Dynamics, German Aerospace Center, Wessling (Munich), Germany.

8. April 23, 1998: "Mobile robot approach and avoidance behaviors with a model of classical and operant conditioning." Invited seminar, Department of Psychobiology, University of Geneva, Switzerland.

9. April 3, 1998: "A neural network that learns approach and avoidance behaviors with a mobile robot." Invited colloquium, Department of Computer Science, Swiss Federal Institute of Technology, Lausanne, Switzerland.

10. March 24, 1998: "Apprendimento, Controllo e Visione: Applicazioni Biologiche, Psicologiche e Tecnologiche della *Dinamica Neurale*" (Learning, control and vision: biological, psychological and technological applications of *Neural Dynamics*). Invited seminar, Istituto San Raffaele, Department of Biotechnology, Milan, Italy.

11. March 23 and 24, 1998: "Introduction to Neural Networks." Four-hour lecture series for undergraduate students. Department of Cognitive Sciences, Istituto San Raffaele, Department of Cognitive Sciences, Milan, Italy.

12. March 17, 1998: " An overview of the Neurobotics Lab: perception, learning and control with mobile robots." Invited colloquium, Center for Neuro-Mimetic Systems, Swiss Federal Institute of Technology, Lausanne, Switzerland.

13. February 17, 1998: "Elaborazione di pattern visivi ed ultrasonici per robot mobili" (Visual and ultrasonic pattern processing for mobile robots). Invited seminar, University of Parma, Italy.

14. February 5, 1998: "Reti neurali non supervisionate per il controllo adattivo di robot mobili" (Unsupervised neural networks for adaptive control of mobile robots). Invited seminar, University of Palermo, Italy.

15. December 10, 1997: "Introduzione alla neurobotica: controllo visivo, localizzazione e riconoscimento di immagini ultrasoniche con un robot mobile" (Introduction to Neurobotics: visual control, localization, and sonar pattern recognition with a mobile robot). Invited seminar, DIST, University of Genova, Italy.

16. December 3, 1997: "Applicazione di una rete neurale per l'apprendimento classico al controllo di un robot mobile" (Mobile robot control with a neural network model of classical and operant conditioning). Invited seminar, DIST, University of Genova, Italy.

17. January 28, 1997: "Adaptation, gain control and retinal processing." Two-hour lecture, Boston University Department of Cognitive and Neural Systems graduate course in vision.

18. November 25, 1996: "Neural networks for adaptive and reactive navigation with wheeled mobile robots." Colloquium in the Department of Computer Science, Colorado State University. Ft. Collins, CO.

19. November 22, 1996: "Computational neural models for mobile robot control and navigation." Colloquium at the Center for Robotics and Intelligent Systems, Colorado School of Mines. Golden, CO.

20. November 21, 1996: "Application of biological neural networks to mobile robot control and navigation." Colloquium at the University of Colorado, Department of Computer Science. Boulder, CO.

21. October 31, 1996: "Light and dark adaptation: psychophysics, physiology, and theory." Invited lecture, Boston University CNS Advanced Vision Seminar, Boston, MA.

22. July 1, 1996: "Applying neural networks to the control and navigation of mobile robots." Colloquium at the Swiss Federal Institute of Technology. Lausanne, Switzerland.

23. June 26, 1996: "Neural networks for content-based storage and retrieval of images in online databases." Invited symposium talk, Second Bodensee Symposium on Multimedia & Online. Friedrichshafen, Germany.

Object recognition system using sonar

M. İhsan Ecemiş and Paolo Gaudiano

Boston University Neurobotics Laboratory
Dept. of Cognitive and Neural Systems
677 Beacon St., Boston MA 02215 USA

ABSTRACT

Sonar is used extensively in mobile robotics for obstacle detection, ranging and avoidance. However, these range-finding applications do not exploit the rich information present in sonar echoes. In addition, mobile robots need robust object recognition systems. The ability to "see" with sound has long been an intriguing concept. Certain animals, such as bats and dolphins, are able to recognize the shape and nature of objects and to navigate using ultrasound. This work aims to set up and develop hardware and software components of an object recognition system using ultrasonic sensors of the type commonly found on mobile robots. Results demonstrate that sonar can be used as a low-cost, low-computation sensor for real time object recognition tasks on mobile robots. This system differs from all previous approaches in its simplicity, robustness, speed and low cost.

Keywords: Object recognition, sonar, robots, ARTMAP.

1. INTRODUCTION

Recently, our group introduced a novel system for recognizing objects using the information extracted from sonar echoes.[1,2] The system we presented uses readily available and inexpensive hardware. Our work was based on the observation that animals such as bats and dolphins can perform remarkable sensory feats using ultrasound signals.[3,4] In contrast, typical robotics applications only use sonar as a range finder, measuring the time-of-flight of the leading edge of the ultrasonic echo to determine the distance to the object that reflected the echo.[5,6]

In our first study we used a Fuzzy ARTMAP neural network[7] to classify echoes from five objects placed at various distances from the sonar. We chose ARTMAP because of its speed, its ability to learn incrementally and its proven performance on a variety of real-world pattern recognition problems. For a description of Fuzzy ARTMAP please refer to the original publication[7] or our earlier article.[1]

Our initial results were very encouraging: the recognition system was able to perform with an accuracy as high as 96%.[1,8] The power spectral density (PSD) was then used as input to the Fuzzy ARTMAP neural network. Later we increased processing speed and flexibility in the data pre-processing scheme and tested an alternative way of extracting information from each echo: the envelope of the echo in the time domain rather than its frequency content.[2,9] In this article we present a systematic study of the system that confirms our previous findings about the value of sonar as a sensor for object recognition.

The remainder of this article is organized as follows: Section 2 describes the system, including the hardware and software components. Section 3 describes the results. The article closes with a short discussion of the results.

2. DATA COLLECTION AND PROCESSING

The hardware system consists of an instrument-grade electrostatic Polaroid transducer, a Polaroid ranging circuit board (series #6500), and a data acquisition board that can operate under the LINUX operating system (DAS16-M1, Computer Boards, Inc., with a LINUX driver written by Warren Jasper of North Carolina State University).

For all results reported here, the sonar module was placed on a movable cart, at approximately the same height as the stand upon which each object was placed. Figure 1(a) shows the sonar setup and a 1-gallon plastic water bottle as a target object. The distance to each object was measured manually and later confirmed directly from the sonar echoes.

Further author information: M.İ.E.: E-mail: ecemis@cns.bu.edu; P.G.: E-mail: gaudiano@cns.bu.edu; Web site: http://neurobotics.bu.edu.
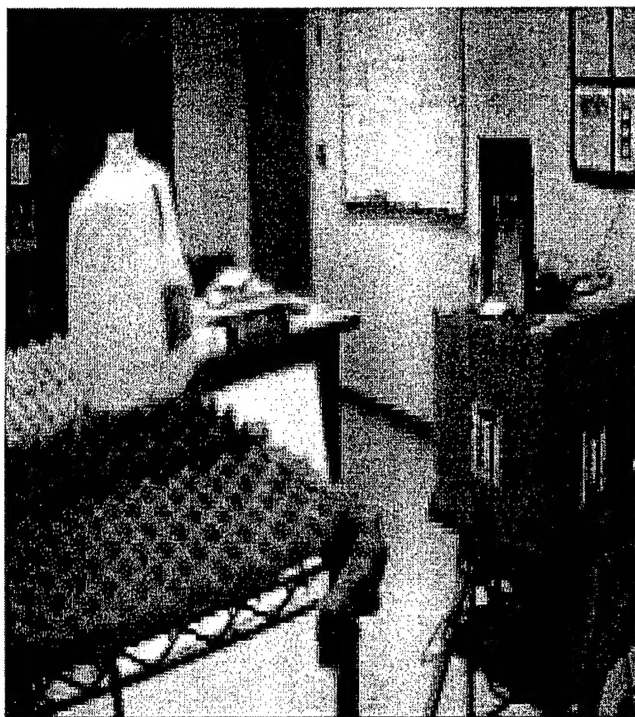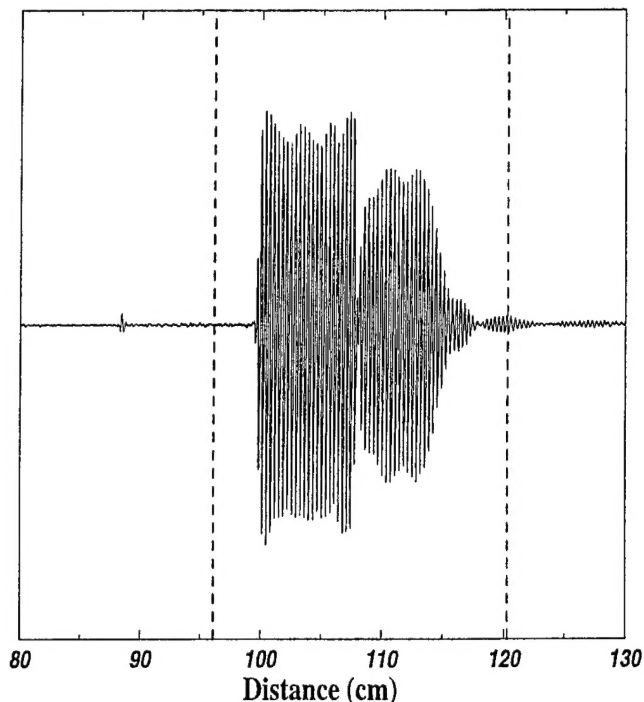
Figure 1. (a) The sonar setup in front of the stand upon which a 1-gallon water bottle is placed. (b) Typical sonar echo after application of the digital bandpass filter. The vertical dashed lines demarcate the portion of the echo used for calculating the PSD function, as described in the text.

Each echo is sampled at a frequency of 500kHz. The raw echo is processed with a digital bandpass filter with cutoff frequencies at 15kHz and 93kHz. The filter is implemented by calculating the inverse Fast Fourier Transform (FFT) of an ideal frequency response (with an amplitude of 1.0 in the desired range and 0.0 elsewhere). A Hamming window is then applied to the inverse FFT to restrict the temporal extent of the filter's impulse response. The amplitude of the impulse response is normalized and truncated to 255 points. This impulse response is then convolved with the echo.

Lowpass or bandpass filters are commonly used in analog-to-digital conversion processes in order to eliminate unwanted frequency components in the source signal. In the current application, the filter reduces noise present in the signal before the echo arrives. This makes the detection of the echo (identification of the exact onset time) highly reproducible; with filtering the system achieved an accuracy of 1 data point (0.002msec). By removing the low-frequency components of the echo, the resulting signal is insensitive to 60Hz line noise and to overall fluctuations that occur for example when the battery that triggers the sonar is running low.

Figure 1(b) shows a sample filtered echo returned by a 1-gal plastic bottle located 100cm in front of the sonar. The vertical lines demarcate the data extracted for calculation of the PSD function, as described below. For the envelope extraction, the data are truncated closer to the echo onset, as described later.

2.1. Extracting Frequency Information

For the results presented in this work, the PSD is calculated using the method of Welch, an approach that combines averaging and windowing. Specifically, we used 18 Hamming windows of 256 points each, with an overlap of 90% between adjacent windows, covering a total of 698 points, which corresponds to 1.4msec of data— or to approximately 24cm in round-trip distance in space. The PSD is calculated by summing the FFT of all 18 windows and dividing the total by 18.

There is a trade-off between the resolution of the estimate and its accuracy. Using a larger FFT window decreases the width of the frequency bins, increasing the resolution of the transform. On the other hand, a smaller FFT window
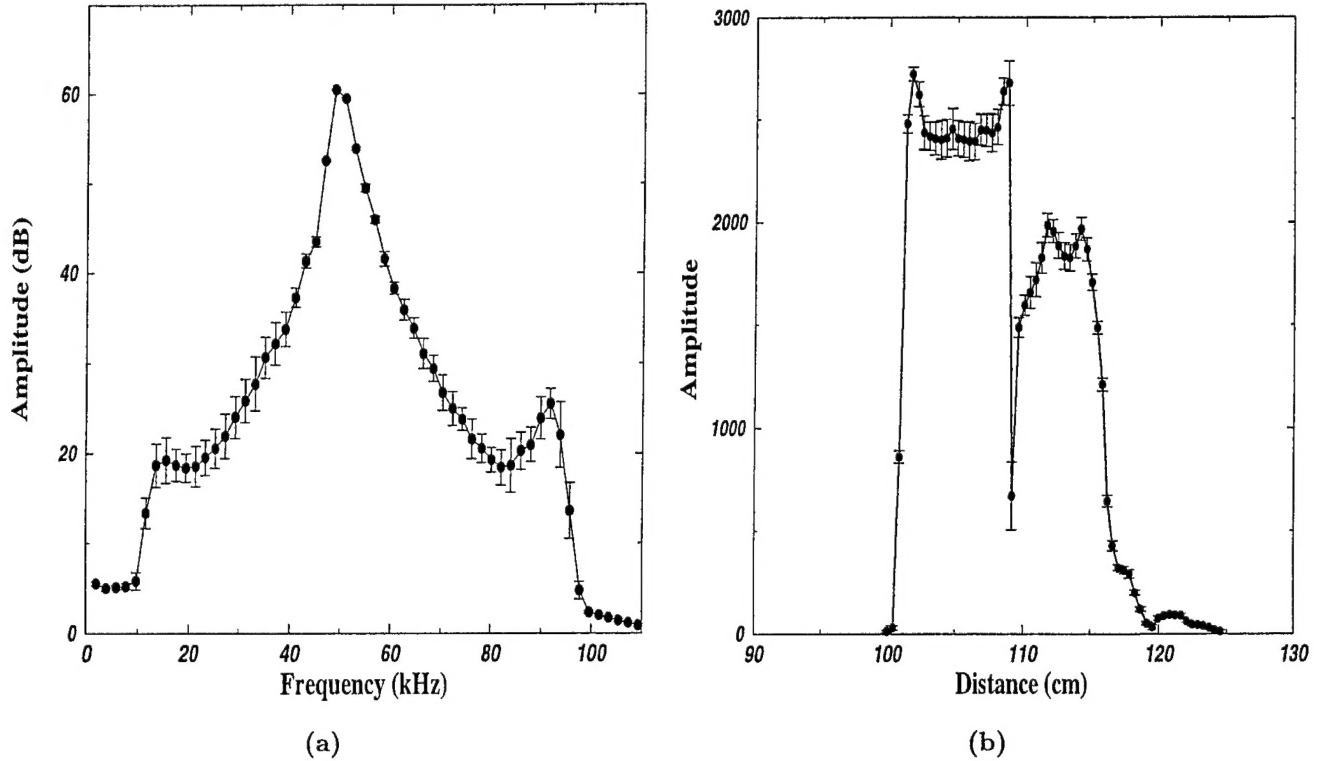
Figure 2. Average (a) PSD and (b) envelope function calculated from 50 echos reflected by the 1-gallon plastic water bottle at a distance of 100cm from the sonar. Error bars show the variance in the 50 measurements.

yields larger frequency bins, effectively averaging nearby bins. The outcome is lower spectral resolution and higher energy accuracy. 256-point Hamming windows are found to be adequate for object recognition purposes. The choice of 90% overlapping windows (instead of the standard 50% overlap) was not crucial for the system performance.

The 256-point Hamming windows yield a resolution of 500kHz/256=1,953Hz. Each PSD vector is truncated to the 40 elements in the frequency range [15,625Hz–91,797Hz], reflecting the characteristics of the band-pass filter. The 40-D vector is used as the input vector for the classifier.

The solid line in Fig. 2(a) illustrates the average PSD obtained from the echo of the 1-gal plastic bottle located at a distance of 100cm from the sonar. Each point along the solid line is the average of 50 measurements (note that 18 Hamming windows are averaged for each of these measurements). The variance in the 50 measurements is shown by error bars. This averaging was done to test the repeatability of the PSD function for a given object at a given distance and aspect. Our results demonstrate that PSD functions are highly repeatable for the same object at the same distance and aspect (notice the low variance in Fig. 2(a)), though they can change significantly as the object is moved or rotated relative to the sonar (not shown here).

2.2. Envelope Function

In this work, we extend our prior research by using time-domain information as input to the classifier, and comparing the recognition performance of the time-domain and frequency-domain methods. In particular, we decided to use the shape of the echo waveform (amplitude modulation of the incoming signal) as input to the classifier for the time domain analysis. The high precision with which we can locate the onset of the echo (1 data point accuracy, corresponding to 2×10^{-6} sec) suggested that shifting in the time domain would not be a problem. If the onset of the echo can not be calculated precisely, the envelope function may not be distance invariant unless used in conjunction with other cues.

Each envelope function is calculated as follows: Starting 30 data points (corresponding to 0.06msec) prior to the echo onset, the code finds the maximum value in each of 60 non-overlapping and contiguous windows of 12 data points each. Since the sampling rate is 10 times higher than the carrier frequency, the processed echo reaches a local

maximum approximately every 10 data points. Twelve data points are found to represent the envelope function adequately. This procedure is very fast and effectively performs down-sampling and half-wave rectification of the original waveform.

Figure 2(b) shows the average and variance of 50 echoes from the 1-gal plastic bottle located at a distance of 100cm from the sonar. The low variance demonstrates that envelope functions are highly repeatable for the same object at the same distance and aspect, though they can change significantly as the object is moved or rotated relative to the sonar.

The entire envelope function consists of 60 points spanning 720/500kHz=1.44msec, which corresponds to a spatial range of about 25cm from the front edge of the object. The envelope function is passed as a 60-D input vector to the classifier. Please note that distance information is not passed to the neural network implicitly or explicitly: all inputs from a given object are classified to the same output node. As described in the next section, classification is always better with the envelope than with the PSD.

3. RESULTS

We performed two experiments to determine the accuracy that can be achieved in recognizing an object independent of its distance from the sonar. Performance of the two methods described in the previous section are compared in these experiments.

3.1. Distance Generalization Experiment

The goal of the first experiment is to show that the system is able to recognize objects using ultrasonic echoes. The tests not only measure system performance when recognizing objects at distances presented during training, but determine how well the system is able to generalize recognizing objects at distances not seen during training.

We collected echoes from four different objects: a 1-liter plastic water bottle, a metal trash can, a styrofoam sheet measuring approximately 34x63cm, and a lego wall measuring approximately 38x10cm. For each object we collected 50 echoes at 11 distances ranging from 50cm to 150cm in 10cm increments. In addition we collected 50 "distractor" echoes from two other objects (a 1-gallon plastic water bottle and a cardboard box) at 100cm only. If there were no distractors, a classifier which randomly selects an object during testing could have an accuracy of 100/4=25%. The distractors decrease "chance performance" further to 100/6≈16.7%.

First, the fuzzy ARTMAP neural network was trained using a randomly selected subset of five processed echoes from each of the four main objects at 90cm, plus five randomly selected echoes from the two "distractor" objects at 100cm. For each training input vector, the desired output class was set to one of six nodes to indicate which object was the correct response.

Learning was set to a single epoch (fast learning mode) with a vigilance level of 0.95. Because ARTMAP is sometimes sensitive to the order of input presentation in fast learning mode, we repeated each experiment 10 times (each time drawing a different random set of 5 processed echoes for each object) and report the average results. However, we found that there was little variance across individual experiments.

Testing was performed only for the four main objects, using the remaining 45 echoes from the distance of 90cm and all 50 echoes from all other distances, i.e., 50, 60, 70, 80, 100, 110, 120, 130, 140 and 150cm. Ignoring the distractors, the neural network was thus trained with only approximately 0.9% of the data. Note that none of the training echoes was used for testing.

Figure 3(a) shows percent accuracy (across all four objects) as a function of distance for this experiment. The solid line with circles shows the average results of 10 ARTMAPs using the 60-D envelope function as input, while the dashed line with squares shows the average results using the 40-D PSD function as input. The error bars in Fig. 3(a) represent the variance in the 10 experiments. Several points merit discussion.

It is clear that the network trained with the envelope function performs better than the network trained with the PSD at all distances. Both schemes work best in the vicinity of the trained distance (90cm) and their accuracy degrades gradually when the object is moved away to "unseen" ranges. The envelope function method yields 100% accuracy even at 100cm even though the network was not trained with any echoes at that distance. This kind of distance dependence in performance is expected because echoes change dramatically with the distance between the sonar and the object that reflected the ultrasound.
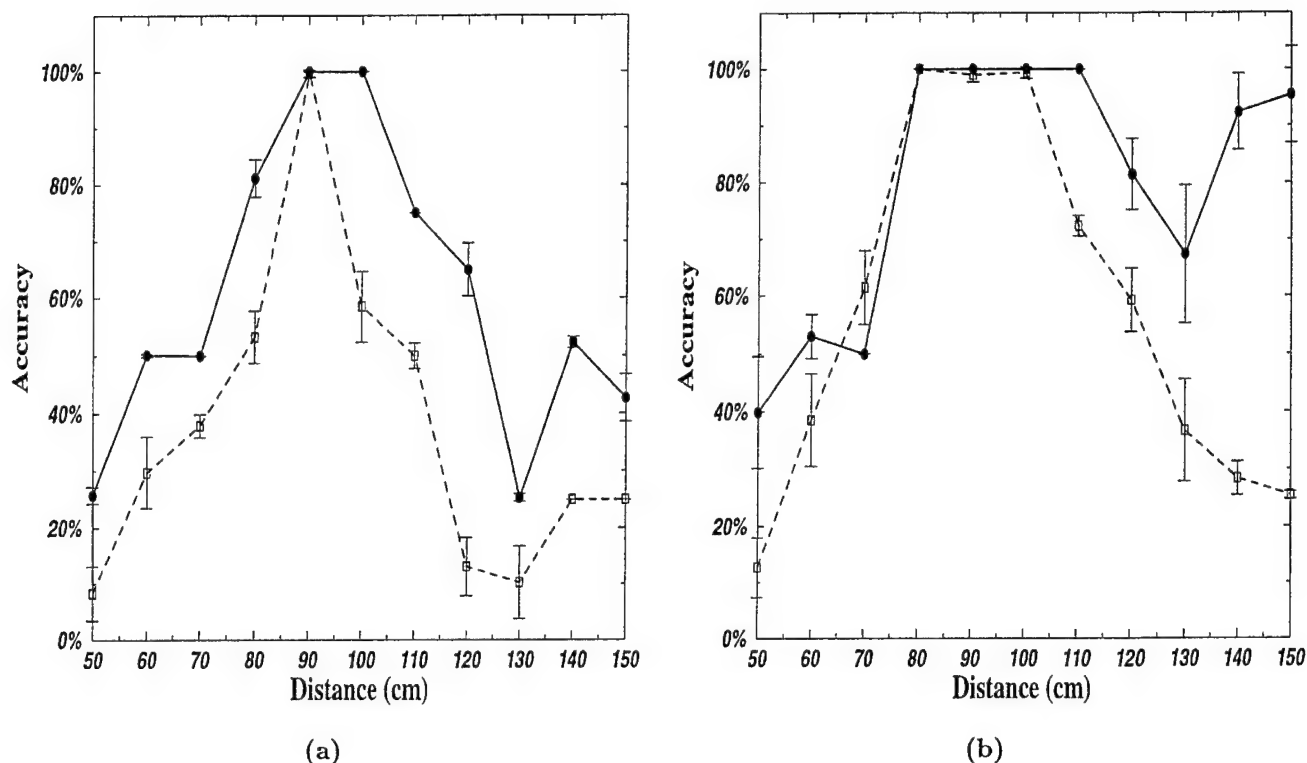
Figure 3. Average recognition accuracy as a function of distance using the envelope (solid line with circles) or PSD (dashed line with squares) as input vectors. **(a)** The network is trained only at 90cm. **(b)** The network is trained at 80, 90 and 100cm.

Next, ARTMAP was trained using a randomly selected subset of five processed echoes from each of the four main objects at distances of 80, 90, and 100cm, plus five randomly selected echoes from the two "distractor" objects at 100cm. Learning occurred in a single epoch (fast learning mode) with a vigilance level of 0.95.

Testing was performed only for the four main objects, using the remaining 45 echoes from the distances of 80, 90 and 100cm, and all 50 echoes from the distances of 50, 60, 70, 110, 120, 130, 140 and 150cm. Thus, ignoring the distractors, the neural network was trained with only approximately 2.7% of the data. None of the training data was used for testing.

Figure 3(b) shows the results of this experiment in terms of percent accuracy (across all four objects) as a function of distance. The solid line with circles shows the average results of 10 ARTMAPs using the envelope function as input, while the dashed line with squares shows the average results using the PSD function as input. The error bars represent the variance in the 10 experiments.

The envelope function method performs better than the PSD method at all distances except at 70cm. As in the first experiment, the envelope function method has an accuracy of 100% at an unseen distance (110cm). The performance of both schemes was equal or better at all distances compared to the previous case when the network was trained only at 90cm. The results show that training at 80 and 100cm increased the accuracy of the network at novel distances (e.g., 130cm) even though no training data from this distance was provided to the network. In summary, one should collect echoes at different distances to improve distance generalization.

Finally, ARTMAP was trained using a randomly selected subset of five processed echoes from each of the four main objects at distances of 50, 70, 90, 110, 130 and 150cm, plus five randomly selected echoes from the two "distractor" objects at 100cm. Learning again occurred in a single epoch (fast learning mode) with a vigilance level of 0.95. Each experiment was repeated 10 times (each time drawing a different random set of five processed echoes for each object) and the the average results are reported.

Testing was performed only for the four main objects, using the remaining 45 echoes from the distances of 50, 70, 90, 110, 130 and 150cm, and all 50 echoes from the distances of 60, 80, 100, 120 and 140cm (the neural network
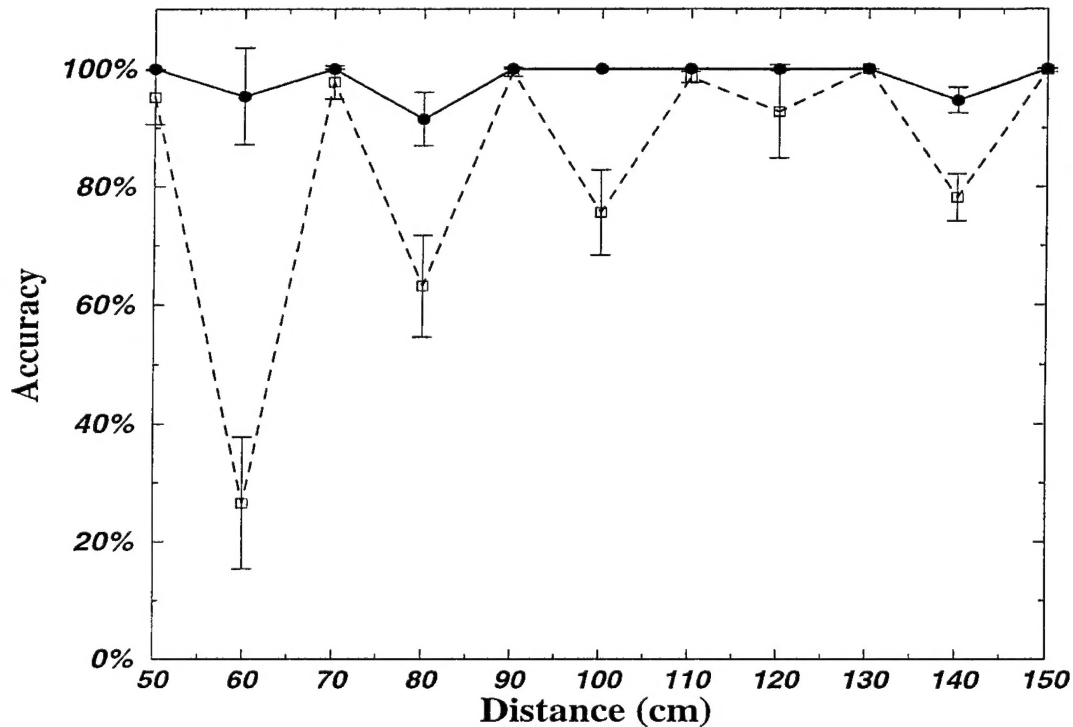
Figure 4. Average recognition accuracy as a function of distance using the envelope function (solid line with circles) or the PSD (dashed line with squares) as input vectors. The network is trained at 50, 70, 90, 110, 130 and 150cm.

was thus trained with only approximately 5.5% of the data). None of the training echoes was used for testing.

Figure 4 shows the results of this experiment in terms of percent accuracy (across all four objects) as a function of distance. The solid line with circles shows the average results of 10 ARTMAPs using the 60-D envelope function as input, while the dashed line with squares shows the average results using the 40-D PSD function as input. The error bars represent the variance in the 10 experiments. Several points merit discussion.

First of all, the envelope function method yields better results than the PSD method once again, consistent with results of the previous experiments. The envelope function method yields 100% recognition at all the distances on which it was trained, as well as for two of the distances on which it was not trained (100cm and 120cm). Performance was at or above 90% at all tested distances. The PSD function also yields 100% accuracy at those distances on which it was trained, but performance is considerably worse at untrained distances.

The performance of both schemes improved at all distances (seen and unseen) compared to the previous cases when the network was trained at only 90cm or at 80, 90 and 100cm. Another interesting observation is that both the envelope function and the PSD methods yield better results and smaller untrained distances. It has been verified informally that this peculiar behavior is the result of an automatic gain mechanism of the Polaroid ranging module, which increases the gain of the receiver in several discrete steps over time to overcome the dissipation of the ultrasonic wave as it travels through the air.

These results confirm the assertion that ultrasonic echoes provide some information about the objects from which they are reflected. From a practical point of view, this suggests that, using the envelope function, it is sufficient to train the object recognition system every 20cm or so, with only a few returns at each distance. With a sonar firing every 100msec, this means that a robot approaching an object head-on can quickly learn to classify the object.

3.2. Object Recognition at Varying Aspects

One important restriction of the results in the previous section is that they are based on the object being "viewed" from a single angle. Given the directional nature of acoustic waveforms, one can expect dramatic changes in the
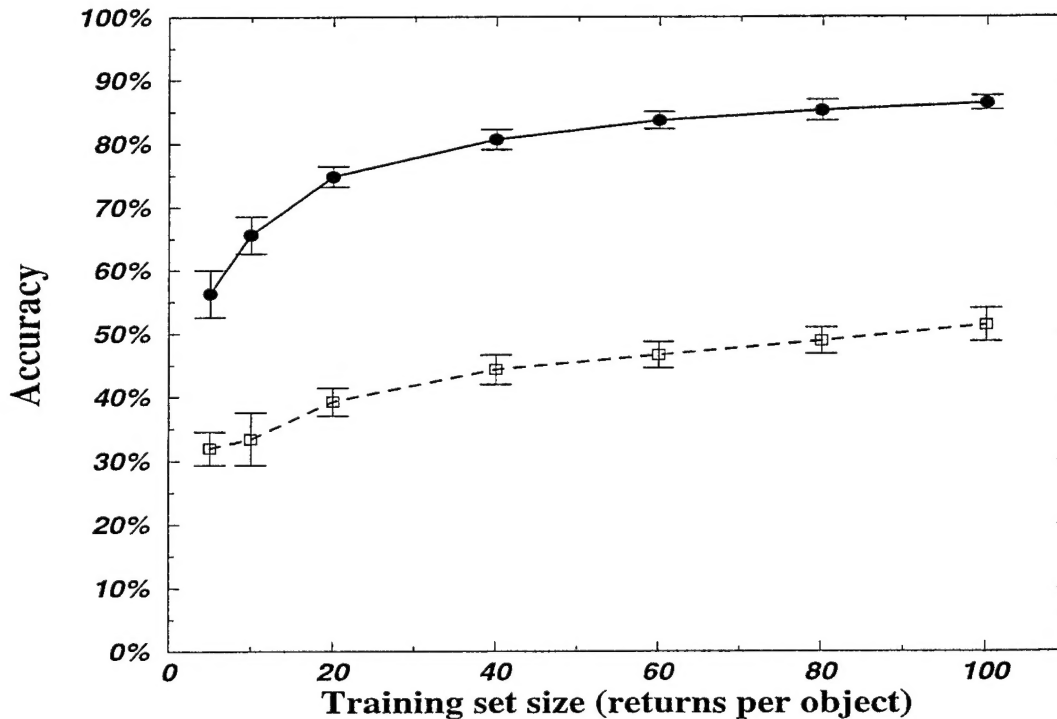
Figure 5. Recognition accuracy as a function of training set size for all five objects using the envelope (solid line with circles) or PSD (dashed line with square) function as input.

PSD and envelope functions as objects are rotated by different angles. Informal observations showed that rotations as little as 5 deg appreciably change the echoes.

In this article we tested object recognition under a fairly unconstrained configuration meant to imitate what might happen with a mobile robot. The sonar was mounted on a rolling cart and could thus be moved relative to each object. For this experiment we used five objects: a 1-gal plastic bottle, a cardboard box measuring 15x25x20cm, a styrofoam sheet measuring approximately 34x63cm, a lego wall measuring approximately 38x10cm, and a bucket-like box (the lego bucket) measuring approximately 18x18x25cm.

For each object, at the start of data collection the cart was moved back-and-forth across angles of +/- 45-deg and distances ranging between about 75cm and 130cm from the object. Because of the weight of the cart, the wheel configuration, and human error, the sonar was not always pointing directly at the object. Nevertheless, this kind of "noise" was left in the data sets. This experiment was meant to replicate a scenario in which a robot is moving around an object.

Data collection for each object lasted about 20 seconds, collecting one sonar echo every 100msec or so, for a total of 200 echoes for each object. During this time the cart was moved from one extreme to the other approximately 5 times.

Training was performed using anywhere between 5 and 100 randomly chosen echoes (out of a total of 200) for each object. The purpose of this experiment was to determine the least number of "views" the system needed to sample in order to recognize objects reliably in their next encounter. Testing only used the echoes that were not seen during training. ARTMAP vigilance was set to 0.9, training lasted one epoch (fast learning), and each experiment was repeated ten times with ten different random seeds. Again, the results were very stable across experiments.

Figure 5 shows recognition accuracy as a function of training set size (out of 200). As before, the envelope function results are shown as a solid line, and PSD results with a dashed line, while the variances are shown by the error bars. Two points are important. First, the envelope function clearly outperforms the PSD, in most cases nearly doubling the accuracy of recognition. Second, even in this relatively unconstrained case, the classifier performs remarkably

well, achieving an overall recognition accuracy of over 55% with only 5 training vectors per object (2.5% of the data set), and nearly 90% accuracy with 100 training vectors per object (50% of the data set).

It is also interesting to consider the efficiency of the Fuzzy ARTMAP neural network. When 40 envelope vectors per object are used for training (20% of the data) and vigilance is set to 0.9, ARTMAP creates a total of 110 category nodes to classify all five objects at all distances and angles, achieving an overall accuracy of 82%. We also tried varying the vigilance parameter but found the results to vary only slightly.

4. CONCLUSIONS

We have presented two simple experiments to show that the sonar system described in this work is able to recognize objects using ultrasonic echoes. The first experiment dealt with the distance generalization capabilities of the system. ARTMAP was trained at certain distances and its performance at untrained distances was examined. The second experiment imitated a robot moving around an object so that the sonar system could sample different aspects of the objects. The envelope function method outperformed the PSD method in both experiments.

The results strengthen the claim that sonar can be used as an effective sensor for object recognition. Our system can easily work in real time, making it possible to sample, process and classify each echo several times per second. Clearly, there are many ways in which we could try to improve our results, for instance by adjusting the pre-processing scheme, using a different classification method, or collecting larger data sets. We could also increase efficiency by using dedicated hardware for some of the pre-processing. However, we feel that the simplicity and robustness of this system are part of its appeal.

This is not the first proposal for the use of sonar for object recognition tasks. Kleeman and Kuc[10] used a sonar array for classification of multiple targets into four reflector types (planes, corners, edges, and unknown), by combining the ranging information from two transmitters and two receivers. Canhui and Regtien[11] used three piezoelectric transducers to distinguish different objects which have the same area of reflection based on comparing the acoustic characteristics of the returns.

Sillitoe et al.[12] used a radial basis function neural network to recognize corners, poles, and other shapes typically found in indoor environment using a *bistatic* sonar array (a bistatic sonar is one in which the transmitter is separate from one or more receivers). Sobral et al.[13] performed recognition of simple objects by generating transfer functions or impulse responses from the envelope of the sonar echo of each of four objects at eight orientations, then using regression to find the best matching object for a given novel input vector (i.e., an unknown object).

Harper and McKerrow[14] used a continuously transmitted frequency modulated (CTFM) sonar. Echoes collected by a second transducer were demodulated with the transmitted signal to produce audio tones proportional to the target range. They used an artificial neural network to classify the return spectra in order to recognize plants. All of these approaches, however, utilize specialized hardware and are thus not easy to replicate.

Our goal eventually is to migrate the entire system on-board one of our robots. So far we have worked with a stand-alone sonar because of our frequent need to modify the setup, but all the components could easily be placed inside any robot with sonar sensors and an on-board PC. We have undertaken such a project in collaboration with IS Robotics (Somerville, MA; www.isr.com).

In the future, it might be necessary to combine this system with other sensor systems such as vision and laser range finders. In the long term, the intention is to develop a system that allows the robot to recognize several objects in arbitrary locations while moving autonomously through an unstructured environment.

ACKNOWLEDGMENTS

REFERENCES

1. W. W. Streilein, P. Gaudiano, and G. A. Carpenter, "A neural network for object recognition through sonar on a mobile robot," in *Proceedings of the International Symposium on Computational Intelligence in Robotics and Automation*, pp. 271–276, 1998.

2. M. I. Ecemiş and P. Gaudiano, "Object recognition with ultrasonic sensors," in *Proceedings of the International Symposium on Computational Intelligence in Robotics and Automation, CIRA'99*, pp. 250–255, 1999.

3. I. E. Dror, M. Zagaeski, D. Rios, and C. F. Moss, "Neural network sonar as a perceptual modality for robotics," in *Neural Systems for Robotics*, O. Omidvar and P. van der Smagt, eds., pp. 1–16, Academic Press, San Diego, 1993.

4. P. W. B. Moore, H. L. Roitblat, R. H. Penner, and P. E. Nachtigall, "Recognizing succcessive dolphin echoes with an integrator gateway network," *Neural Networks* **4**, pp. 701–709, 1991.

5. J. Borenstein, H. Everett, and L. Feng, *Navigating mobile robots*, A.K. Peters, Wellesley, MA, 1996.

6. J. J. Leonard and H. F. Durrant-Whyte, *Directed sonar sensing for mobile robot navigation*, The Kluwer international series in engineering and computer science, Kluwer Academic Publishers, Boston, 1992.

7. G. A. Carpenter, S. Grossberg, N. Markuzon, J. H. Reynolds, and D. B. Rosen, "Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps," *IEEE Transactions on Neural Networks* **3**(5), pp. 698–712, 1992.

8. W. W. Streilein, *Neural Networks for Classification and Familiarity Discrimination, with Radar and Sonar Applications*. PhD thesis, Boston University, 1998. Available through UMI, Ann Arbor, MI.

9. M. I. Ecemiş, *Development of a Sonar-Based Object Recognition System*. PhD thesis, Boston University, 2000. Available through UMI, Ann Arbor, MI.

10. L. Kleeman and R. Kuc, "Mobile robot sonar for target localization and classification," *The International Journal of Robotics Research* **14**, pp. 295–318, 1995.

11. C. Canhui and P. P. L. Regtien, "A smart sonar object recognition system for robots," *Measurement Science and Technology* **4**(1), pp. 95–100, 1993.

12. I. Sillitoe, D. J. Mulvaney, and A. Visioli, "The classification of ultrasonic signals using multistage radial basis function neural networks," in *Proceedings of the IASTED International Conference on Signal and Image Processing*, N. M. Namazi, ed., pp. 135–139, IASTED, ACTA Press, (Anaheim, CA, USA), 1995.

13. L. Sobral, R. Johansson, G. Lindstedt, and G. Olsson, "Ultrasonic detection in robotic environments," in *Proceedings of the 1996 IEEE International Conference on Robotics and Automation*, vol. 1, pp. 347–352, IEEE Press, 1996.

14. N. L. Harper and P. J. McKerrow, "Recognition of plants with ctfm ultrasonic range data using a neural network," in *Proceedings of the 1997 IEEE International Conference on Robotics and Automation, Albuquerque, NM, USA*, pp. 3244–3249, 1997.